Telehealth Ecosystems in Practice
M. Giacomini et al. (Eds.)

doi:10.3233/SHTI230793

Subjective Cognitive Load Evaluation of a Mobile Personal Health Record Application

Neşe ZAYİM[a], Hasibe YILDIZ[a,1] and Yılmaz Kemal YÜCE[b]
[a]*Department of Biostatistics and Medical Informatics, Akdeniz University, Turkey*
[b]*Department of Computer Engineering, Alanya Alaaddin Keykubat University, Turkey*
ORCID ID: Neşe ZAYİM https://orcid.org/0000-0003-3133-2692
Hasibe YILDIZ https://orcid.org/0000-0003-1971-5164
Yılmaz Kemal YÜCE https://orcid.org/0000-0003-4023-0401

Abstract. Mobile Personal Health Records (mPHRs), which make it possible to track and manage users' health information, can be an important aid in improving people's health. Despite its potential benefits, poor usability of systems can hinder the adoption and use of mPHRs. This study aims to evaluate the usability of a mobile health application in terms of perceived cognitive workload and performance. The cognitive workload experienced by 30 volunteers (15 experienced and 15 inexperienced), was measured while performing the given tasks with the NASA-Task Load Index (NASA-RTLX) scale, and the duration of the fulfillment of the tasks by eye tracking device. While there was no significant difference between the two user groups in the completion time of the tasks, a significant difference was found in the perceived cognitive load. "Making an appointment", which could take much longer to complete than other tasks, resulted in the highest cognitive load for all users. Further usability research using think-aloud protocols and user interviews could provide insights into design improvements for reducing cognitive load and enhancing performance.

Keywords. Personal Health Records, mPHR usability, cognitive load, performance

1. Introduction

Personal Health Records (PHR), make it possible to access, enter, manage and share users' health information in a private, secure and confidential environment [1], integrate external information systems such as laboratory, pharmacy and healthcare provider systems [2, 3], and are systems intended to be easy to access and use [4,5]. Despite the many prospective benefits and potential improved health outcomes associated with mPHR and the features they may offer, the main challenge for adoption and usage of mPHR is poor system usability [6-8].

Usability can be broadly defined as the capacity to allow users to perform their tasks safely, effectively, efficiently and with pleasure [9,10]. HIMMS (Healthcare Information and Management Systems Society) [11] covers the nine basic principles of usability; "simplicity, naturalness, consistency, acceptability and feedback, effective use of

[1] Corresponding Author: Hasibe YILDIZ, E-mail: yildizhasibe@hotmail.com, Phone: +90 (551) 154 4160.

language, efficient interactions, effective information presentation, preservation of content and reduction of cognitive load".

Considering the limited capacity of the human mind, cognitive load theory [12] suggests optimizing the presentation and design of information to improve mental performance.

Because a person starts out with a very limited pool of cognitive resources, a poor design can easily deplete him. As a user's cognitive load increases, the performance of the user decreases, and the probability of making mistakes increases rapidly [13].

While PHRs were initially web-based applications, the number of mobile PHRs (mPHRs) applications is increasing rapidly with the penetration of mobile platforms into every aspect of our daily lives. The widespread use of mobile devices has led to the rapid launch of mPHRs applications in all major mobile markets, making it easier for consumers to access these applications [14]. Compared to desktop devices, mobile devices have features to be considered in design such as smaller screen, device use with finger gestures and virtual keyboard, lower processing power, the need to work on battery for a long time, and limited data entry. Through careful evaluation of mobile devices and user-friendly interface design, cognitive demands on users can be minimized [15].

The aim of this preliminary study is to examine the effect of user experience on performance and perceived cognitive workload based on the e-Nabız [16] - a mobile personal health record application offered to the citizens by the Ministry of Health of Turkey.

2. Methods

A convenience sample of 30 adult e-Nabız users participated in this study. Inclusion criteria were that volunteer participants aged 18-59, using smartphones with Android platform operating system and using e-Nabız app as a novice or expert.

To determine the cognitive workload that occurs during the use of e-Nabız mPHR application, 30 volunteer users, were asked to perform 9 tasks given in scenarios through the application. Half of the users who are expert users (EU) had been using the application for more than 1 year, the remaining 15 inexperienced or with little experience novice users (NU) had been using the application for less than a year. Written informed consent was obtained from all participants.

The tasks to be fulfilled are: T1:"making an appointment, T2:"canceling an appointment", T3:"reviewing physician visit information on a specific date", T4:"adding reminder information to the prescribed medication from the last doctor visit", T5:"entering allergy information", T6:"reading the package insert of the most recently prescribed drug", T7:"entering the medication side effects", T8:"Reviewing information about a diagnosed condition" and T9:"viewing the most recent medical certificate".

After each task was completed, the subjective cognitive workload was measured with the NASA-RTLX questionnaire (9 NASA-RTLX scores per participant, a total of 135 NASA-RTLX scores for each group). Users' task completion time in milliseconds was recorded for each task using the eye tracker Pupil Core.

NASA-TLX was developed by Hart and Staveland [7] consisting of the subsections "Mental Demand", "Physical Demand", "Temporal Demand", "Performance", "Effort" and "Frustration", which must be scored for each task within a scale of 0-100 point range in increments of 5 points. In addition, the subsections must be compared in pairs regarding their perceived importance. Byers et al. [6] simplified the model and removed

the comparison. This modified model, called NASA-RTLX still allows for a high experimental validity [8]. The overall cognitive workload that the participant experiences is determined through an addition of the scores, and then division on the six different dimensions to get the average. We then calculated the difference between experienced users (EU) and novice users (NU) in the 95% confidence interval using the Mann-Whitney U test.

3. Results

From the participants, shown in Table 1, 17 users have high school or lower qualifications, 13 users have college or higher degrees. 26 of the participants had been using smartphones more than 5 years. Most of them had been using Android platform six years or more. 15 persons had experience with the e-Nabız mPHR more than a year. Users mostly used the application to make appointments, examine the laboratory results and radiology reports, and drug checking.

Table 1. Demographic information of the recruited user

Demographic Information		Frequency	%
Education	<=High School	17	56.7
	>High School	13	46.3
Smartphone usage	1-5 Years	4	13.3
	>=6 Years	26	86.7
Android usage	1-5 years	8	26.7
	>=6	22	73.3
mPHR usage	Never	4	13.3
	Occasionally	13	43.3
	1-2 times/month	11	36.7
	At least once a week	2	6.7
Web or/and App	None	4	13.3
	App	13	43.3
	Web	2	6.7
	Web+app	11	36.7
e-Nabız experience	Never used	6	20.0
	< 1 Year	9	30.0
	1-2 Years	14	46.7
	>= 3 Years	1	3.3
Used e-Nabız for:			
Make an appointment		26	86.7
View lab results		17	56.7
view radiology reports		12	40.0
Drug checking		10	33.3
others		6	20.0

The comparison of the NASA-RTLX scores with the task completion times of each task can be seen in Table 2. In both evaluation methods, the average values of all users per each task were calculated. NASA-RTLX score averages range from 0-100, and as cognitive load increases, the score increases.

Table 2. NASA-RTLX and Task Time Mean Scores by Tasks

	NASA-RTLX			Task Time		
Task	Mean	95% Confidence Interval for Mean(s)	Std. Deviation	Mean	95% Confidence Interval for Mean(s)	Std. Deviation
T1	22.81	16.33 to 29.28	17.34	138.72	107.9 to 169. 53	82.52
T2	15.5	9.44 to 21.56	16.23	28.26	21.75 to 34.77	17.44
T3	21.25	14.97 to 27.53	16.81	44.23	35.94 to 52.52	22.19
T4	20.64	14.76 to 26.52	15.75	52.64	43.45 to 61.83	24.62

T5	18.17	12.92 to 23.41	14.05	47.69	36.65 to 58.73	29.57
T6	15.14	9.29 to 20.99	15.68	38.01	29.11 to 46.91	23.84
T7	19.31	12.85 to 25.76	17.3	40.12	27.80 to 52.44	32.99
T8	19.53	13.71 to 25.35	15.58	32.38	27.58 to 37.17	12.84
T9	10.69	6.84 to 14.54	10.31	25.69	16.66 to 34.73	24.19

By examining participants' experience, although there are no significant differences between users' task completion time (Table 3), we found that novice users (151.46), defined as those have experience less than one year, had significantly higher cognitive workload ($p<0.001$). Except physical dimension, novice users also had significantly higher scores on mental ($p=0.001$), temporal ($p=0.012$), performance ($p<0.001$), effort ($p<0.001$) and frustration ($p=0.024$) dimensions of cognitive load.

Table 3. Mann-Whitney U test of differences between experienced users (EU) and novice users (NU)

	User experience	Mean Rank	Mann-Whitney U	p-value
Mental	EU	119.54	6957.500	0.001
	NU	151.46		
Physical	EU	132.70	8734.000	0.544
	NU	138.30		
Temporal	EU	123.73	7524.000	0.012
	NU	147.27		
Performance	EU	115.52	6415.500	<0.001
	NU	155.48		
Effort	EU	118.67	6840.500	<0.001
	NU	152.33		
Frustration	EU	125.09	7706.500	0.024
	NU	145.91		
NASA-RTLX	EU	118.63	6835.000	<0.001
	NU	152.37		
TaskTime	EU	134.47	8973.500	0.828
	NU	136.53		

4. Discussion

In this study, new users and expert user groups consisting of each 15 people were formed to determine performance and the cognitive load arising from the use of the e-Nabız mPHR application offered to the citizens in Turkey.

The research shows that there is no significant difference in user performance between the groups EU and NU. The tasks completed in the shortest time are T8 and T2. The task "making an appointment", which takes much longer time compared to other tasks, at the same time, the perceived cognitive load is highest in this task, because many steps had to be performed to accomplish this task. Simplifying and facilitating appointment scheduling will help the application to be adopted more easily.

NASA-RTLX averages obtained range from 10.69 to 22.81 on a scale of 0-100. Tasks with the lowest perceived cognitive load are Task9 and Task2. The reason why the cognitive load is not very high may be that these two tasks are relatively simple, can be accomplished in a short time, and do not require any data entry. If the given tasks contain information that needs to be remembered and the design is not descriptive enough, the user perceives a high cognitive load and has difficulty in fulfilling the tasks [17]. Based on these results, we can predict that more difficult and comprehensive tasks may cause higher cognitive load in novice users.

5. Conclusion

Keeping in mind that PHRs target a broad range of users with different demographics, such as smart phone skills, education and health literacy etc., it is crucial to conform usability criteria in the design of these system [18]. Findings of this study limited to explain user differences in terms of performance and cognitive load, doesn't provide information about design issues related to cognitive load and performance. Detailed analysis of usability data from think-aloud protocol, which is a common method in usability testing, where users think out loud while performing tasks and user interviews of the ongoing research is expected to shed light on these issues in mobile personal health record application.

References

[1] Foundation M. The Personal Health Working Group: Final Report. Markle Foundation [Internet]. 2003 July 1. Available from: https://markle.org/publications/1429-personal-health-working-group-final-report/.
[2] Jones DA, Shipman JP, Plaut DA, Selden CR. Characteristics of personal health records: findings of the Medical Library Association/National Library of Medicine Joint Electronic Personal Health Record Task Force. J Med Libr Assoc. 2010;98(3):243-9. doi: https://doi.org/10.3163/1536-5050.98.3.013. PubMed PMID: 20648259; PubMed Central PMCID: PMC2900995.
[3] Showell C. Barriers to the use of personal health records by patients: a structured review. PeerJ. 2017;5:e3268.
[4] Clarke MA, Schuetzler RM, Windle JR, Pachunka E, Fruhling A. Usability and cognitive load in the design of a personal health record. Health Policy and Technology. 2020;9(2):218-24. doi: https://doi.org/10.1016/j.hlpt.2019.10.002.
[5] Wang T, Dolezel D. Usability of web-based personal health records: an analysis of consumers' perspectives. Perspectives in Health Information Management. 2016;13(Spring).
[6] Byers JC, Bittner A, Hill SG. Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary. Advances in industrial ergonomics and safety. 1989;1:481-5.
[7] Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Advances in psychology. 52: Elsevier; 1988. p. 139-83.
[8] Bustamante EA, Spain RD, editors. Measurement invariance of the NASA-TLX. Proceedings of the human factors and ergonomics society annual meeting; 2008: SAGE Publications Sage CA: Los Angeles, CA.
[9] Henderson A. Interaction design: beyond human-computer interaction. Ubiquity. 2002;2002(March):2.
[10] Preece J, Rogers Y, Sharp H, Benyon D, Holland S, Carey T. Human-computer interaction: Addison-Wesley Longman Ltd.; 1994.
[11] HIMMS. 2022. p. https://www.himss.org/resources/nine-essential-principles-software-usability-emrs.
[12] Sweller J. Cognitive load theory. Psychology of learning and motivation. 55: Elsevier; 2011. p. 37-76.
[13] Tracy JP, Albers MJ, editors. Measuring cognitive load to test the usability of web sites. Annual Conference-society for technical communication; 2006.
[14] Kharrazi H, Chisholm R, VanNasdale D, Thompson B. Mobile personal health records: an evaluation of features and functionality. Int J Med Inform. 2012;81(9):579-93. Epub 20120717. doi: https://doi.org/10.1016/j.ijmedinf.2012.04.007. PubMed PMID: 22809779.
[15] Alshehri F, Freeman M, Freeman AE. Are you smart enough for your smart phone? A cognitive load comparison. 2013.
[16] e-Nabız - Personal Health System: Turkish Ministry of Health; [cited 2023 30.05.2023]. Available from: https://enabiz.gov.tr/.
[17] Pachunka E, Windle J, Schuetzler R, Fruhling A. Natural-setting PHR usability evaluation using the NASA TLX to measure cognitive load of patients. 2019.
[18] Taha J, Sharit J, Czaja SJ, editors. Usability of an electronic personal health record (PHR) among a diverse group of adults. Proceedings of the human factors and ergonomics society annual meeting; 2014: SAGE Publications Sage CA: Los Angeles, CA.